

A disembodied developmental robotic agent called Samu Bátfai

Norbert Bátfai,*

November 11, 2015

Abstract

The agent program, called Samu, is an experiment to build a disembodied DevRob (Developmental Robotics) chatter bot that can talk in a natural language like humans do. One of the main design feature is that Samu can be interacted with using only a character terminal. This is important not only for practical aspects of Turing test or Loebner prize, but also for the study of basic principles of Developmental Robotics. Our purpose is to create a rapid prototype of Q-learning with neural network approximators for Samu. We sketch out the early stages of the development process of this prototype, where Samu's task is to predict the next sentence of tales or conversations. The basic objective of this paper is to reach the same results using reinforcement learning with general function approximators that can be achieved by using the classical Q lookup table on small input samples. The paper is closed by an experiment that shows a significant improvement in Samu's learning when using LZW tree to narrow the number of possible Q-actions.

Keywords: *machine intelligence, developmental robotics, deep Q-learning, Liv-Zempel-Welch tree.*

1 Introduction

At first glance, it seems impossible to develop a natural language chatter bot that can understand conversations as humans do, and that is only based on the verbal content of the speech because human communication is only 7 percent verbal [5, pp. 26]. Nevertheless we have started to develop such a system, called Samu. Or, to be more precise, it is called Samu Bátfai due to it will be taught primary in a family circle where the author's three children, ages 7 and 9 years old will also be partially its caregivers.

One of the main design feature of agent program Samu is that we can communicate with it through only one written channel. For example, a consequence of this is that we cannot create associations between a visual channel and an auditory channel. We must work only from the verbal content, but this constraint is what allows us to be able to purely focus on the investigation of basic

*N. Bátfai is with the Department of Information Technology, University of Debrecen, H-4010 Debrecen PO Box 12, Hungary, e-mail: batfai.norbert@inf.unideb.hu.

principles of Developmental Robotics [29]. And there is a further practical reason why Samu has only limited input capability. The character channel will be controlled and monitored very well by caregivers. It is not of great importance at this moment but it will be much more important in the future. Because if we think about it, we are not to permit either our little children or our little DevRob agents to use internet without supervision.

This paper is organized as follows. First, it introduces the architecture of Samu. Then it explains how Samu is linked to Developmental Robotics in a conceptual level. Finally we evaluate our results supplemented with further conceptual elements.

2 Samu

The verification principle [29] is the most important principle of Developmental Robotics. At the beginning of the developmental process, program of Samu will not possess the ability to verify the knowledge. Mostly this is because, like young children of age 2-4 years old, young Samu accepts what its caregivers say without reservation. That is why family caregivers have their own dedicated communication channels on standard input/output. Using these channels, Samu will be able to distinguish its family caregivers, as it is shown in the next lines

```
Norbi@Caregiver> ___next caregiver
Nandi@Caregiver> I am Nandi
```

where the message started with the prefix `___` (three underscores) are not sent directly to Samu. Talking on the internet with other conversationalists and especially with robots that are similar to Samu will be available at a later development stage. In both of these stages, the verification will be able to be carried out only on the basis of Samu's accumulated knowledge and experience. It seems obvious that the knowledge acquisition of a developmental robotic agent can be programmed in a natural way by using some reinforcement learning method. The general architecture of a system of this kind is shown in the following section.

2.1 An idealized architecture

The paper [24] gave the idea for the idealized architecture of Samu that is shown in Fig. 1(a) where Samu reads the sentences from a children's tale or from a conversation. Then it tries to reconstruct the meaning of the read sentences in the form of a 2D simulation game generated from the sentences. Like in paper [24], the screenshots of the generated game will be processed by a Q-learning network, but whereas there the scores are part of the game, here they are based on the comparison of the reality of the reconstructed game and the reality of the input sentences. Theoretically, it would be a hybrid architecture that combines a model-free (Q-learning) and a knowledge-based (generating the 2D programs) approaches. To clarify matters, this architecture would use a model-free method to develop knowledge-based models. But because worth of every model is what we can implement from it, we will strongly simplify it in the next sections.

2.2 Simplified architectures

Starting from the idealized architecture shown in Fig. 1(a) we base the comparison of the reality and the simulated one derived from the input sentences on the principle of consciousness oriented programming (the agents must predict the future and the conscious ones can really foresee and predict it [6]). It is shown in Fig. 1(b) where the simulation program's task is to predict the next input sentence. This modified architecture works with SPO triplets (Subject-Predicate-Object, [27]) rather than the raw sentences being read. Fig. 1(c) shows a further simplified architecture. In this case, we do not have to run the simulation program to predict the next input sentence, instead the prediction is done by the Q-learning component itself. The architectural components of Fig. 1 are discussed in detail in the next points.

2.3 The first rapid prototype

We use a hybrid architecture of the latter two sketches in Fig. 1 for implementing the first rapid prototype of Samu. In the following, this architecture is referred to as *Samu's cognitive engine*.

2.3.1 NLP

As we mentioned earlier, SPO triplets [27] are used instead of sentences. For example, if Samu reads the sentence "*The little mouse has become a big lion.*" the NLP (Natural Language Processing) component should produce the following triplets: i) "*(mouse, become, lion)*" ii) "*(mouse, is, little)*" iii) "*(lion, is, big)*". Link Grammar [1] is used to obtain SPO triplets but automatically identifying these is a complex data mining problem. At this moment we are not focusing on this matter. We are going to partially implement some algorithms of [23] but for this paper, we use only a very simple algorithm shown in code snippet in Fig. 2.

2.3.2 Visual imagery

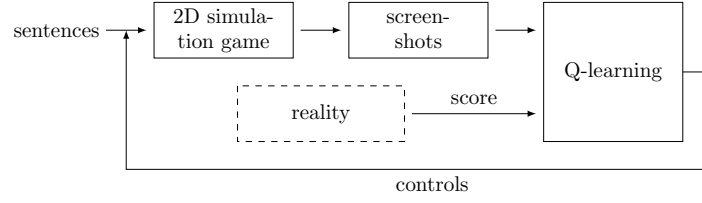
To describe the simulation programs we use a very simplified object oriented style language, where statements have the following form:

`Subject.Predicate(Object);`

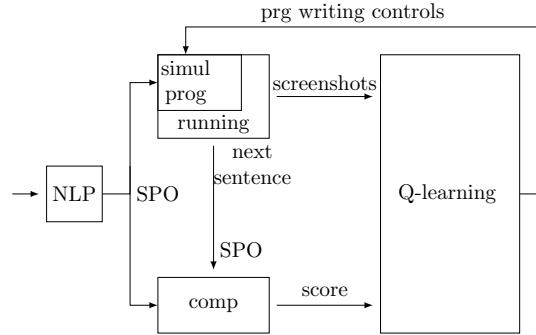
where each identified triplet corresponds to a statement such as this. A simulation program is a statement list that only contains a given number or less statements. In addition, we write it into an image because it would be difficult to handle the statement list as a textual list of statements as it is shown in Fig. 1(c) where the input of the Q-learning is already an image. And finally, it may be noted that the sequence of such "mental" images can be interpreted as visual imagination of Samu. Fig. 3 shows a sample mental image.

2.3.3 Q-learning

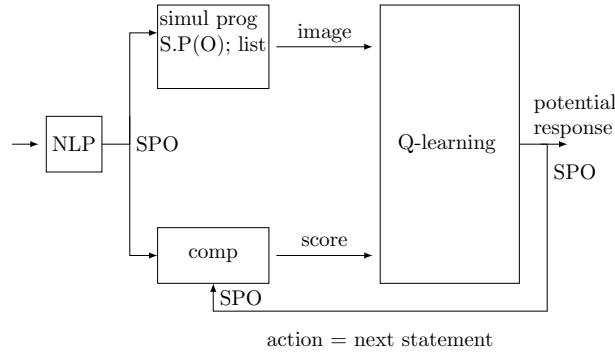
Sigmoidal multilayer perceptron (MLP) neural networks are used to approximate the Q-function. In accordance with the QCON model [22], [30] each



(a) An idealized architecture for planning Samu. As far as it will be possible, the comparison of the reality of the reconstructed game and the reality of the input sentences will have to be automated.



(b) This modified architecture enables fully automated comparison between the reality and the simulation. The score represents the positive reward if the previous predicted (simulated) sentence matches the actual input string. And a negative reward is given if the predicted (simulated) sentence of the previous step differs from the actual input string.



(c) This is a simplified version of the previous architecture where the next sentence is determined by the action of Q-learning. And notice that the simulation program is not being run. Instead, the state of the Q-learning is only the "source code" itself. It is a very simplified simulation program that consists only of statements that have the following form: Subject.Predicate(Object).

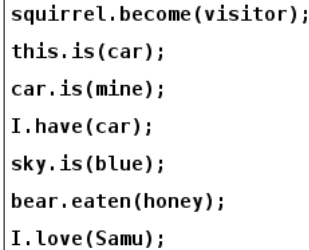
Figure 1: This figure shows some architectural sketches that build on each other to help refining the initial concept of Samu.

```

for ( int k {0}; k<linkage_get_num_links ( linkage ); ++k ) {
    const char* c = linkage_get_link_label ( linkage, k );
    if ( *c == 'S' ) {
        triplet.p = linkage_get_word ( linkage, k );
        alter_p = words[linkage_get_link_rword ( linkage, k )];
        triplet.s = words[linkage_get_link_lword ( linkage, k )];
    }
    if ( *c == 'O' ) {
        triplet.o = words[linkage_get_link_rword ( linkage, k )];
        if ( triplet.p == words[linkage_get_link_lword ( linkage, k )] ) {
            triplets.push_back ( triplet );
        } else if ( alter_p == words[linkage_get_link_lword ( linkage, k )] ) {
            triplet.p = alter_p;
            triplets.push_back ( triplet );
        }
    }
}
}

```

Figure 2: This is a simplified code snippet that is iterated over all linkages of the parsed sentence returned by the Link Grammar [1]. The 'words' is a vector of strings that contains all words of a given linkage. The complete source code in which this snippet is incorporated can be found in the file `nlp.cpp`. This transitional solution can handle the simplest sentences like "A rare black squirrel has become a regular visitor to a suburban garden" , "This is a car" , "This car is mine" , "I have a little car" , "The sky is blue" , "The little brown bear has eaten all of the honey" , "I love Samu". It may be noted that the first one is the example sentence used in the paper [27].



```

squirrel.become(visitor);
this.is(car);
car.is(mine);
I.have(car);
sky.is(blue);
bear.eaten(honey);
I.love(Samu);

```

Figure 3: This is a sample mental image (256x256 of size) that contains sentences shown in the caption of Fig. 2.

possible action has an unique neural network to approximate the Q-value associated with the given state. Each MLP contains three layers with 65536 input units, 80 hidden layer units and an output neuron which gives the approximation of the Q-value of the state-action pair. The states are "mental" images introduced in the previous paragraph. Several examples can be found in the literature of using neural networks to approximate the Q-function, for example, see the book [20] or the tutorial [21] by Gosavi. Alg. 1 shows the used algorithm to perform the Q-learning. This pseudo-code is a simplified and generalized version of the most precise description of the used algorithm that can be found in

the source file `ql.hpp` at <https://github.com/nbatfai/samu/blob/master/ql.hpp>. The regular part of the algorithm is taken from the book [26, pp. 844].

Algorithm 1 Q-learning-with-NN-approximation (s', t') **returns** a triplet (an action)

Input: s', t' $\triangleright s'$ is the current state and t' is the current triplet.
 r', a', p, nn, q \triangleright Local variables: r' is the current reward, a' is the current action and p is a perceptron, $nn, q \in \mathbb{R}$.
 $s, r = -\infty, a, N_{sa}, Prcps_a$ \triangleright Persistent variables: s is the previous state, r is the previous reward, a is the previous action, N_{sa} is the state-action frequency table and $Prcps_a$ is the array of the perceptrons.

Output: a' $\triangleright a'$ is an action, the predicted next triplet.

```

1: procedure SamuQNN( $s', t'$ )
2:    $r' = compare(t', a)$ 
3:    $a' = t'$ 
4:   if  $r > -\infty$  then
5:     Increment  $N_{sa}[s, a]$ 
6:      $nn = Prcps_a[a](s)$ 
7:      $q = nn + \alpha(N_{sa}[s, a])(r' + \gamma \max_p Prcps_a[p](s') - nn)$ 
8:     Back-propagation-learning( $Prcps_a[a], s, q, nn$ )  $\triangleright$  where  $q - nn$  is the error
      of the MLP  $Prcps_a[a]$  for input  $s$ .
9:      $a' = argmax_p f(Prcps_a[p](s'), N_{sa}[s', p])$ 
10:  end if
11:   $s = s'$ 
12:   $r = r'$ 
13:   $a = a'$ 
14:  return  $a'$ 
15: end procedure

```

2.4 Conceptual background of Samu

In order to outline the conceptual background we should begin with the explanation of the usage of the notion of disembodiment. It is simply an indication that Samu project is *100 percent pure software*. We can say that Samu is a robot without a body. Certainly, this does not mean a return to GOFAI. All the more so, because we are going to follow the recommendations given as replies to the critics of Dreyfus and Dreyfus [26, pp. 1025].

It is certainly arguable whether the principle of embodiment is met or not. We believe that we are not confronted with this principle as it will also be shown in the following.

Returning to the basic principles formulated by [29], there are two different use cases in respect of the validation processes of *Samu's cognitive engine*. If we observe how Samu operates we can easily distinguish a short-term verification from a long-term one. It is clear that the short-term verification of Samu's knowledge is trivial because it is the prediction itself. This type of verification is a quintessence of the whole system. But the long-term verification is a much harder issue since this is only a naive intuitive expectation. The long-term verification should give the answer to the question: How can we extend the short-term verification for longer time periods. In this sense the long-term verification means the understanding the conversations and read texts. It is

an open question how much the distance between these two kinds of verifying the knowledge. Based on our experiences we have used a quantity called *bogo-relevance* that may be positioned somewhere in the middle of these two extremes of verification. Using the notations of Alg. 1 the computation of the *bogo-relevance* of a prediction q in the state s can be rewritten as follows:

$$\begin{aligned} \text{bogo_relevance}(s, q) &= \\ &= \frac{\text{Prcps}_a[q](s) - \overline{\text{Prcps}_a}}{\max_p \text{Prcps}_a[p](s) - \min_p \text{Prcps}_a[p](s)} \end{aligned}$$

where the $\overline{\text{Prcps}_a}$ is the mean of the computed values of the perceptrons that are applied to the same given input state s , that is

$$\overline{\text{Prcps}_a} = \frac{\sum_{p \in \text{Prcps}_a} \text{Prcps}_a[p](s)}{|\text{Prcps}_a|}$$

here $|\text{Prcps}_a|$ denotes the number of all perceptrons in the structure Prcps_a . The program prints the values $100 * \text{bogo_relevance}(s, q)$ that can be seen in Fig. 6.

From a spiritual point of view, Samu's soul consisting of the weights of Samu's ANNs. Because they are only simple numbers (and not quantum qubits, for example), Samu's soul can be saved to a file. This file is called `samu.soul.txt` and it is created automatically when Samu's program receives a given (SIGINT, SIGTERM, SIGKILL or SIGHUP) signal. But since the soul cannot live without body, Samu's body is the UNIX process that runs the source code of Samu. The actions of Samu are his statements and replies that are formed by the close interaction with the external world. In the light of this we think that, we are not confronted with the spirit of the principle of embodiment [29].

The fulfillment of the principle of subjectivity [29] is trivially met since it follows from using human caregivers to train Samu. Accordingly, the a priori subjectivity of caregivers implies the subjectivity of their robots.

It is easy to see that the criteria of the principle of grounding [29] have also been fulfilled during the short-term verification. Because the set of possible outcomes {good prediction, bad prediction} may be assigned to each action, since each action is a predicted sentence in our model.

Finally, it may be noted that in the sense of the terminology introduced in [6], Samu may become a conscious or even an intuitive computer program since it will be able to predict the future better than a random guess.

3 Results

The results and evaluations are divided into three experiments. In the first experiment, we verify Samu's implementation of Q-learning by comparison with using the classical Q lookup table. Here we have only used a small number of sentences. The second experiment extends the investigation to a larger sample. Finally, the third experiment has already focused only on the learning of the larger corpus where using the Liv-Zempel-Welch (LZW) [31] dictionary tree to narrow the scope of selecting Q-actions seems promising for accelerating Samu's learning.

3.1 Experiment 1: "SAMU A"

Our main purpose in this paper is to equip Samu with the ability to carry out reinforcement learning with general function approximators. To validate this we use the following experiment. The sentences shown in the caption of Fig. 2 are considered as a short story and Samu is being trained continuously with this during the experiment. Samu reads the sentences of the story and predicts the next story sentence. A third of a point is given if a triplet member is predicted correctly. The reward is based on the comparison of triplets and is computed by the following equation

```
double reward =
    3.0 * triplet.cmp ( prev_action ) - 1.5;
```

therefore 10.5 points are given if all predictions are correct and conversely - 10.5 points are given if all predictions are wrong. We implement both the simple Q lookup table and the general function approximators based solutions in order to compare them. The learning curves for these two experiments can be seen on Fig. 4 and Fig. 5. For example, the first one shows that Samu can perfectly memorize the text (consisting of the seven test sentences) after he read it roughly sixty times. It can be seen well on Fig. 5 that the general function

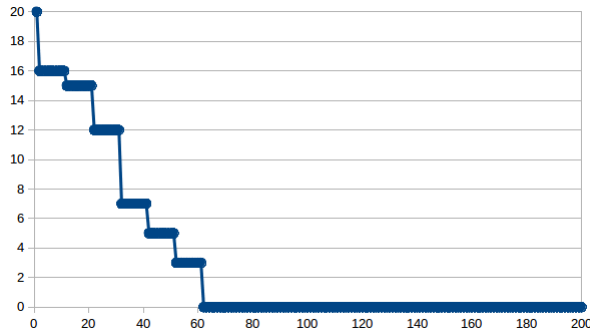


Figure 4: The horizontal axis shows the number of trials using Q lookup table. Trials consist of reading the 7 given sentences. The vertical axis shows the prediction error of the next sentences.

approximators based solution can also learn the true predictions. So we may assume that our neural network based approximation works well. Therefore, in the following we are not going to use the Q lookup table either in this paper or in the source codes. (Lookup table-based code snippets can still be found in the source files. These have been wrapped in conditional compilation directives but they are deprecated in the new versions of Samu, see for example [10], [11], [12].)

In order to allow the evaluation and verification of these results, we create a GitHub repository for the source of this study. It is called **samu** and available at <https://github.com/nbatfai/samu>. The two use cases of Samu shown in Fig. 4 and Fig. 5 are presented in a YouTube video at the following link <https://youtu.be/qAhxVQk-dvY>. It is our hope that Samu will be the ancestor of DevRob chatter bots that will be able to chat in natural language like humans do. However, it should be noted that in order to use Samu as a chat program

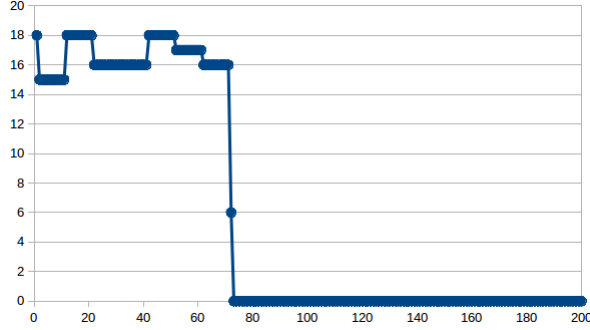


Figure 5: The horizontal axis shows the number of trials using general function approximators. The vertical axis is the same as in the case of Fig. 4.

we need to try another fork of the project Samu [15]. The first official forks, called Isaac [10], Jacob [11], Judah [12], Perez [13], Hezron [9], Ram [14] and Amminadab [8] are introduced in next section.

3.1.1 Samu’s family tree

Samu and his descendants (at this moment Isaac, Jacob, Judah, Perez and Hezron) are open-source programs released under the GNU GPL v3 or later. Samu is a generic name of these projects that are intended to investigate deep reinforcement learning methods to facilitate the development of a chat robot that will be able to talk in natural language like a human does. At this moment they are case studies of using Q learning with neural networks for predicting the next sentence of a conversation. They can only be run on Unix-type (such as GNU/Linux) environments.

a) Samu [15] is the initial project to start creating a DevRob chatbot agent. This project is presented in detail in this paper. Samu’s development is frozen to allow the evaluation and verification of the results shown in Fig. 4 and Fig. 5. Therefore this project is continued in the project entitled Isaac. *b)* The project Isaac [10] has been forked from Samu. Isaac takes a step towards the deep Q learning, for example, Isaac’s perceptrons are accelerated by using CUDA and OpenMP. *c)* The project Jacob [11] has been forked from Isaac. Jacob replaces Isaac’s picture-based visual imagination with a character-based one. *d)* The project Judah [12] has been forked from Jacob. Judah equips Jacob with a text-based user interface (TUI) that can be seen in Fig. 6. In addition, Judah can also be run on both remote computers as well as on supercomputers. *e)* The project Perez [13] has been forked from Judah. Using this, we can make experiments to see the behaviour of Samu on different parameter settings of the incremental learning. *f)* The project Hezron [9] has been forked from Perez. In this project we are going to compare the arrangements of SPO triplets in the visual imagery. For example, a pyramid-shaped arrangement can be seen in a YouTube video at <https://youtu.be/zjoINedftPY> or a fully justified one is shown in Fig. 6. *g)* The project Ram [14] has been forked from Hezron. Ram allows to experiment with different deep Q learning algorithms to express feelings when he is talking. *h)* The project Amminadab [8] has been forked from Ram. Amminadab uses the LZW tree to narrow the scope of selecting

```

nbatfai@orchmach: ~/GitHubRepos/judah

Samu's visual imagery —
1. who.are(you);who.are(you);who.are(you);who.are(you);who.are(yo
2. i.am(robot);i.am(robot);i.am(robot);i.am(robot);i.am(robot);i.a
3. what.is(name);what.is(name);what.is(name);what.is(name);what.is(name);what.
4. name.is(Judah);name.is(Judah);name.is(Judah);name.is(Judah);name.is(Judah);
5. you.live(where);you.live(where);you.live(where);you.live(where);you.live(wh
6. I.live(Debrecen);I.live(Debrecen);I.live(Debrecen);I.live(Debrecen);I.live(
7. you.are(how);you.are(how);you.are(how);you.are(how);you.are(how);you.are(ho
8. i.am(one);i.am(one);i.am(one);i.am(one);i.am(one);i.am(one);i.am(one);i.am(
9.
10.

Samu's answers —
Judah@sleep.599.10%> what is name
Judah@sleep.599.10%> name is Judah
Judah@sleep.599.10%> you live where
Judah@sleep.599.9%> I live Debrecen
Judah@sleep.599.9%> you are how
Judah@sleep.599.10%> i am one

Caregiver shell —
Norbi> Type your sentence and press [ENTER]

```

Figure 6: A screenshot of Judah’s TUI. The top window contains the contents of Samu’s visual imagery. The mid window displays Samu’s responses. For example, the prompt `Judah@sleep.599.9%` contains the name of the robot, an at sign (@), the state of the robot, the number of the known triplets and a *bogo-relevance* value. The bottom window is the text input field where the caregiver can enter text and commands for Samu.

Q-actions.

3.2 Experiment 2: "JUDAH A"

Religious robot children As we have already mentioned earlier, in this paper we focus on the shaping of the neural infrastructure of Samu rather than on the triplet extraction. But certainly we need corpora for training and testing Samu. Evidently that we used only a meaningless sequence of sentences in the previous example of the seven test sentences. It was enough to verify the implementation of the used algorithms but it is not suitable for teaching a chat robot. We need real corpora and it is equally important to use sophisticated triplet extraction algorithms that can preprocess these corpora. In addition, from the point of view of Developmental Robotics, the essential constraint of the selection of the training corpora is that these may only contain a limited number of distinct words. An example of such corpus may be the Basic English Bible (BBE, Bible in Basic English) [25] that contains 850 words [2]. We have tried to use the Gospel of Matthew. But it should be noted that the initial simple algorithm shown in Fig. 2 can extract only roughly 700 triplets from it. This is due to the following two reasons. *i)* The used algorithm is elementary in the sense that it tries to apply only a very simple rule to extract triplets as exactly shown in Fig. 2 *ii)* In the raw text source of BBE, the ‘.’, ‘:’ and ‘;’ letters were replaced by a newline character and the numbers were deleted when we applied the triplet extraction mechanically to each line. (And for example,

there were lines that contains only one word. The results of this preprocessing can be found in the files called `bbe` and `bbe.triplets` that are available at the author's university page <http://shrek.unideb.hu/~nbatfai/>.) It follows that this corpus of the 700 SPOs still has not been regarded as a text that can hold meaning for human readers. In a second experiment, we use this corpus and another small one consisting of the sentences shown in Fig. 7. Because

```
{
  "introduce_myself",
  {
    "Who_are_you",
    "I_am_a_robot",
    "What_is_your_name",
    "My_name_is_Judah",
    "Where_do_you_live",
    "I_live_in_Debrecen",
    "How_old_are_you",
    "I_am_one_year_old",
    "Where_were_you_born"
    "I_was_born_in_Debrecen"
  }
}
```

Figure 7: A small training corpus that has already held meaning for human readers.

of the larger number of all sentences of these two used corpora we give a more sharp comparison for triplets and the rewarding policy. The reward is changed by the following equation

```
double reward =
  ( triplet == prev_action ) ? 1.0 : -2.0;
```

therefore in the case of the introductory text 10 points are given if all "responses" are correct and conversely -20 points are given if all "answers" are incorrect. The size of the input and hidden layers of MLPs is reduced due to we use a character console-based visual imagination (that was introduced by the project Jacob [11]). The neural networks corresponded to the triplets are equipped with 800 input units and 32 hidden layer units in this experiment. Fig. 8, Fig. 9 and Fig. 10 show the learning curve for teaching the smaller (the introductory) text contained in Fig. 7.

In connection with using sacred texts for training Samu's neural networks, the question arises whether religious training should be given to robots? Samu may transform this question an theoretical-ethical one into a practical one. In a more pragmatic sense, it would be an interesting linguistic and theological challenge to attempt to make a new translation of a Gospel into SPO triplets.

Family circle robotics Another possibility to learn Samu's neural networks is to use the conversational interactions of CHILDES (Child Language Data Exchange System) [3]. But we believe that the social interactions with family caregivers will become the mainstream in teaching Samu. For this reason, Samu saves all conversations that take place with the caregivers as training files.

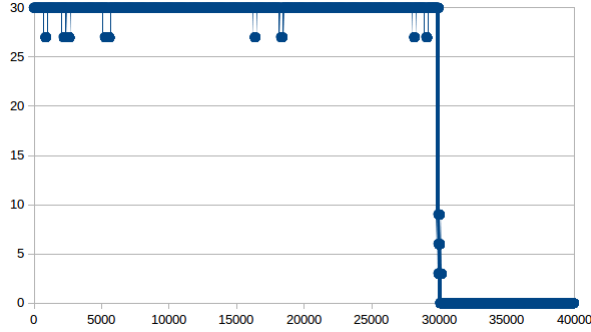


Figure 8: The learning curve for the introductory text contained in Fig. 7. The axes are the same as in Fig. 4 and 5. This graph can be verified and reproduced by using Judah’s git repository [12] tagged as `v0.0.1.exp.a1`. The execution takes a couple of hours on an ordinary PC. A detailed description of using the program can be found in the Judah project’s wiki at <https://github.com/nbatfai/judah/wiki>.

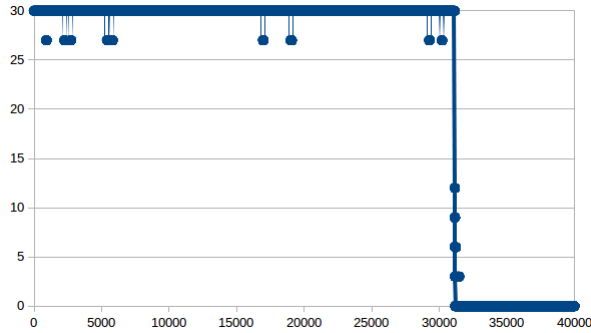


Figure 9: This graph is the same as the previous one except that before the measurement, Samu has been trained with the bigger corpus for several thousand times. The axes are the same as in the previous figure.

3.2.1 Social interaction with each other

Perhaps the most important possibility is that when the Samu-type DevRob agents share their knowledge among each other. At this moment the development of the communication protocol is in progress.

3.3 Experiment 3: "JUDAH and PEREZ B"

This experiment is the same as the previous one, but here Samu learns the larger corpus. It should be noticed that perfectly memorizing a verbatim text of 700 sentences is really difficult or nearly impossible for most humans, as well. Accordingly, the algorithms developed in Judah have not performed well. Fig. 11 shows that Samu (to be more precise Judah) learns very slowly and it may even be that he does not learn. We must try to accelerate the learning. A promising acceleration mechanism is tested with the smaller corpus in Fig. 12.

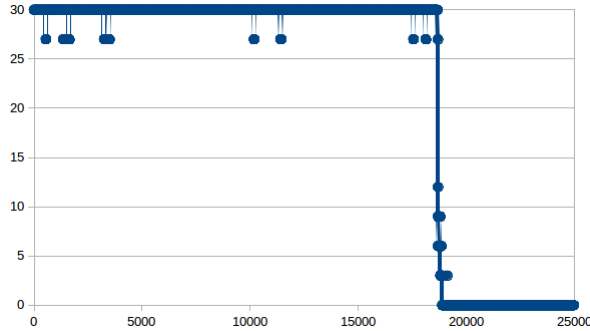


Figure 10: In this figure, the learning curve of Fig. 8 is recomputed with the modified number of tries for triplets in the exploration function. There is a YouTube video at <https://youtu.be/ONgfQDY1qvo> that shows how to use Judah to reproduce this figure.

We have modified the usage of the exploration function. In addition we use a trick of incremental learning [4], [18]. Using a trivial version of this, Samu already can learn the larger corpus. It is shown in Fig. 13 and Fig. 14. In the beginning, we taught only the first seven sentences of the larger corpus as a training text. After Samu had already learnt it, we added the next seven sentences of the larger corpus to the training file. This incremental procedure was repeated continuously during the experiment. However it is clear that it will not be enough if the number of triplets continues to increase. For example, during the experiment, it was well felt that the phenomenon of catastrophic forgetting [18] has slowed the speed of Samu’s learning. This is also reinforced by the fact that the incremental learning was implemented in the `main.cpp` and not in a class of Samu. From a point of view of a programmer, it means that the incremental approach is not yet really part of Samu. We think there may be another possibility to improve the performance of Samu’s neural networks. In our near future work, we are going to carry out experiments to see whether it is possible to increase the speed of learning by improving Samu’s visual imagery. Experimenting with several (such as pyramid or justified) arrangements of triplet statements in the visual imagery does not bring improvements directly. A more interesting possibility would be to add new internal information to the mental images. These information might be interpreted as feelings, for more details in this direction see the project Ram [14].

3.4 Experiment 3: "AMMINADAB B"

With using the LZW dictionary tree to narrow the scope of selecting Q-actions, we can significantly improve Samu’s learning as it is shown in Fig. 15 and Fig. 16. In our case, the nodes of the LZW tree are triplets. For example, the next log snippet shows the tree that has been built at the end of the first incremental step.

```

----2__ son was Hezron
__1__ sons were Zerah
----2__ sons were Zerah
__1__ sons were and

```

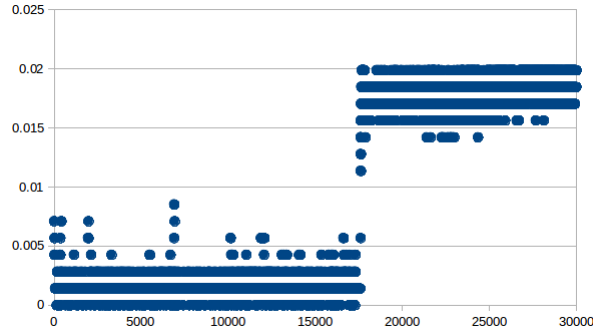


Figure 11: In this figure we have not used the reward-based scale that was used in the previous ones. Now the vertical axis simply shows the quotient of number of good predictions and the sum of the good and the bad ones. For example, the value .02 means roughly 14 good and 690 bad predictions. This experiment has been done with a variant of the Judah version tagged by `v0.0.2.exp.b1` (the name of the BBE corpus must be specified in line 236 (see at <https://github.com/nbatfai/judah/blob/master/main.cpp#L236>). In sum, Samu has not learnt the larger corpus yet in this experiment).

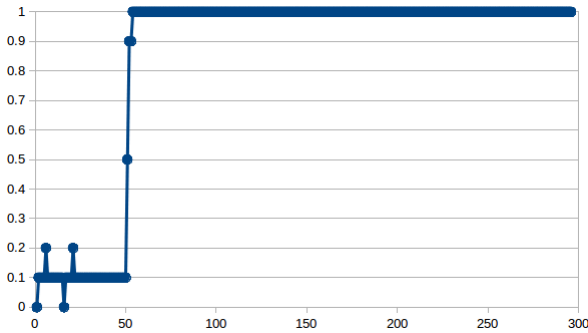


Figure 12: This graph shows that Samu succeeded in learning the smaller corpus. The source code for this experiment can be found on the Judah repository [12] tagged by `v0.0.2.exp.b1` (see also at <https://github.com/nbatfai/judah/releases/tag/v0.0.2.exp.b1>).

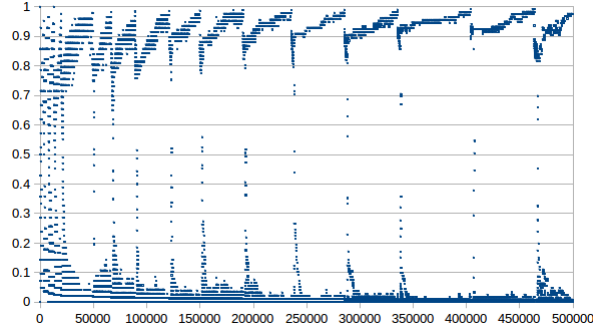


Figure 13: Samu managed to learn the larger corpus now. To be more precise he has learnt the first 126 (18×7) sentences out of roughly 700 (during 18 incremental steps). The source code for this experiment can be found on the Perez repository [13] tagged by `v0.0.1.exp.b1` (see also at <https://github.com/nbatfai/perez/releases/tag/v0.0.1.exp.b1>). For achieving the same learning curve like this the cache file `bbe.triplets` needs to exist in the directory from where Samu was executed.

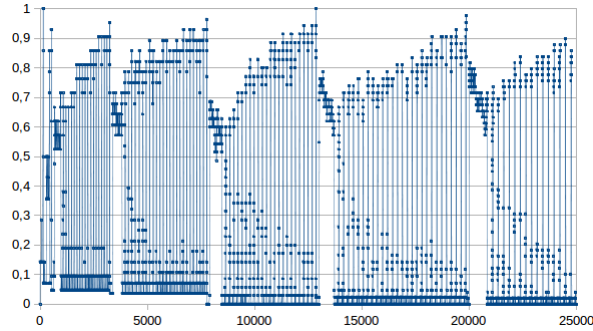


Figure 14: This figure shows only the first 25000 points of the previous one. It is interesting to note that the success of the learning of the incremental training corpora (when the quotient in question is greater than roughly 0.95) is followed by a short stable period in which there is no significant forgetting.

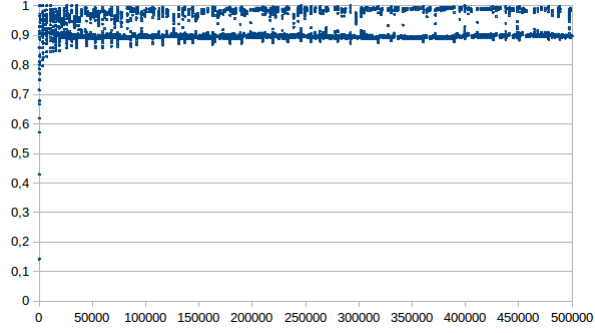


Figure 15: This figure repeats the experiment shown in Fig. 13 by using Amminadab project [8]. It shows very well that Amminadab can significantly improve Samu's learning curve. He has already learnt the first 392 (56×7) sentences out of 704 (during 56 incremental steps).

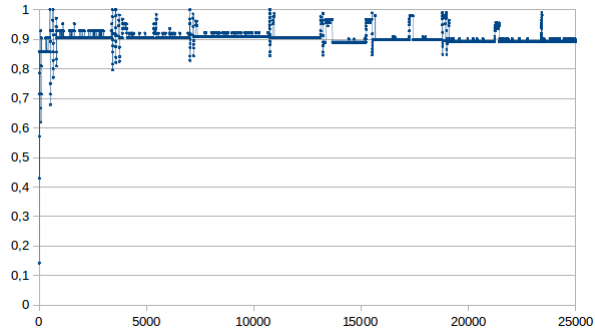


Figure 16: Similar to Fig. 14 this figure shows only the first 25000 points of the previous experiment shown in Fig. 15.


```

__1__ son was Amminadab
-----3__ son was Amminadab
-----2__ son was Ram
__1__ son was Hezron
-----7__ son was Amminadab
-----6__ son was Ram
-----5__ son was Hezron
-----4__ sons were Zerah
-----3__ sons were and
-----2__ son was Jacob
__1__ son was Isaac
__1__ son was Jacob
-----2__ son was Amminadab
__1__ son was Ram
0__

```

where the number in the beginning of each line shows the depth of a given triplet. During our experiments, the maximum depth of the LZW tree is restricted to 10. The modified version of Alg. 1 is shown in Alg. 2.

Algorithm 2 Q-learning-with-LZW-narrowing (s' , t') **returns** a triplet (an action)

Input: s' , t' $\triangleright s'$ is the current state and t' is the current triplet.
 r' , a' , p , nn , q , \mathcal{B} \triangleright Local variables: r' is the current reward, a' is the current action and p is a perceptron, nn , $q \in \mathbb{R}$, \mathcal{B} is a set of triplets.
 s , $r = -\infty$, a , N_{sa} , $Prcps_a$, \mathcal{N} \triangleright Persistent variables: s is the previous state, r is the previous reward, a is the previous action, N_{sa} is the state-action frequency table and $Prcps_a$ is the array of the perceptrons, \mathcal{N} denotes the actual node of the LZW tree.

Output: a' $\triangleright a'$ is an action, the predicted next triplet.

```

1: procedure SamuQNN( $s'$ ,  $t'$ )
2:    $r' = \text{compare}(t', a)$ 
3:    $\mathcal{N} = \text{build\_lzw\_tree}(t')$ 
4:    $a' = t'$ 
5:   if  $r > -\infty$  then
6:     Increment  $N_{sa}[s, a]$ 
7:      $nn = Prcps_a[a](s)$ 
8:      $\mathcal{B} = \{\text{children of the node } \mathcal{N}\}$ 
9:      $q = nn + \alpha(N_{sa}[s, a])(r' + \gamma \max_{p \in \mathcal{B}} Prcps_a[p](s') - nn)$ 
10:    Back-propagation-learning( $Prcps_a[a]$ ,  $s$ ,  $q$ ,  $nn$ )  $\triangleright$  where  $q - nn$  is the error
    of the MLP  $Prcps_a[a]$  for input  $s$ .
11:     $a' = \text{argmax}_{p \in \mathcal{B}} f(Prcps_a[p](s'), N_{sa}[s', p])$ 
12:  end if
13:   $s = s'$ 
14:   $r = r'$ 
15:   $a = a'$ 
16:  return  $a'$ 
17: end procedure

```

Finally, we may remark that Amminadab applies some cellular automaton steps in his visual imagery:

```

for ( int i {1}; i<nrows-1; ++i )
  for ( int j {1}; j<ncols-1; ++j )
    console2[i][j] = console[i-1][j]+console[i][j-1]
                      +console[i+1][j]+console[i][j+1];

```

In addition in this experiment the SARSA [26, pp. 844] version of Q-learning is applied because in this case it is not necessary to use the restriction $p \in \mathcal{B}$ introduced in line 9 of Alg. 2. The precise details can be found in source files `samu.hpp` and `ql.hpp` in the Amminadab’s GitHub repository [8].

4 Summary and highlights of Samu

Samu B  tfai is a disembodied developmental robotic chat software agent that is intended to become the basis of a chat system which will be able to talk and read in natural language like humans do. Samu

- directly implements the definition of machine consciousness (such as the definitions of conscious and intuitive computer programs) introduced in the paper [6].
- corresponds the principles of Developmental Robotics [29].
- applies Q learning with neural networks approximators
- uses multilayer perceptrons for approximation of the Q function, it is a deep learning [24], [17], [19] feature.

5 Conclusion

Now that we have a working prototype we will be able to start deep Q-learning experiments to investigate Samu’s behavior in big data corpora. It is certainly an important problem in this direction is the triplet extraction. We believe that this can be handled well with the existing NLP tools. As we highlighted earlier, we did not focus on this issue moreover we want to keep it apart from the mainstream of Samu development. But since it will be a very important data mining component of Samu we are going to try to develop it in programming competitions. The first competition has already been announced for our university students. For details, see the author’s university page <http://shrek.unideb.hu/~nbatfai/>. In a similar situation we have already used competitions [16] to catalyze our research programs but this area in question has its own challenges like the Turing test or the Loebner prize [28]. We would like to participate in these challenges in the future.

Returning to the investigation of Samu’s behavior in big data corpora, in applying Alg. 1 there is a principal problem that it uses the state as an index of the frequency table at this moment. This is not a viable option for real applications. But on the one hand the frequency table may be dispensed and on the other hand the current triplet also be used as a simplified state in indexing the frequency table. The next critical part of Samu’s code is the big number of possible actions. Working at the level of 2 years old children Samu should use roughly 2000 words. If we calculate only 700 (S) \times 300 (P) \times 700 (O) triplets we have 1.47×10^8 combinations. Let’s remember that each combination corresponds to a perceptron and the memory footprint of a perceptron is about 40 megabytes of memory right now. But it is also true that this is the worst case, because it strongly depends on the settings of the perceptrons and it can be reduced drastically by using Jacob’s character console-based visual imagination.

And fortunately, the distribution of these combinations is Pareto rather than uniform so this aspect also seems to be handled well.

As additional further work we plan to develop an intellectual (fractal) dimension [7] based benchmark to test the intellectual capabilities of Samu type agent programs. In our case the intellectual dimension may be derived naturally from the triplet prediction used by Samu type agents.

In summary, in this paper we have begun to develop rapid prototypes to support experiments using the deep reinforcement learning techniques to facilitate the creation of a chat robot that will be able to talk in natural language like a human does. The generic name of these prototypes is Samu. At this moment Samu has already worked formally as chat program but despite that we thought of him as a child who is in the prenatal development stage. According to the plans Samu will be learning in family circle when he is born. We hope this paper can help you fork your own family chatbot from Samu projects.

Acknowledgment

The computations used for the experiments shown in this paper were partially performed on the NIIF High Performance Computing supercomputer at University of Debrecen. But it is important to note that the results in this paper can be reproduced on an ordinary PC.

References

- [1] opencog/link-grammar, 2015. URL <https://github.com/opencog/link-grammar>.
- [2] Bible in Basic English, 2015. URL <http://www.o-bible.com/dlb.html>.
- [3] CHILDES - Child Language Data Exchange System, 2015. URL <http://childes.psy.cmu.edu/>.
- [4] R. R. Ade and P. R. Deshmukh. Methods for incremental learning: A survey. *Int. Journal of Data Mining & Knowledge Management Process*, 3(4):119–125, 2013.
- [5] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida. Cognitive Developmental Robotics: A Survey. *IEEE T. Autonomous Mental Development*, 1(1):12–34, 2009.
- [6] N. Bátfai. Conscious machines and consciousness oriented programming. *CoRR*, abs/1108.2865, 2011. URL <http://arxiv.org/abs/1108.2865>.
- [7] N. Bátfai. Turing’s Imitation Game has been Improved. *ArXiv e-prints*, abs/1509.00584, 2015. URL <http://arxiv.org/abs/1509.00584>.
- [8] N. Bátfai. Amminadab, 2015. URL <https://github.com/nbatfai/amminadab>.
- [9] N. Bátfai. Hezron, 2015. URL <https://github.com/nbatfai/hezron>.
- [10] N. Bátfai. Isaac, 2015. URL <https://github.com/nbatfai/isaac>.

- [11] N. Bátfai. Jacob, 2015. URL <https://github.com/nbatfai/jacob>.
- [12] N. Bátfai. Judah, 2015. URL <https://github.com/nbatfai/judah>.
- [13] N. Bátfai. Perez, 2015. URL <https://github.com/nbatfai/perez>.
- [14] N. Bátfai. Ram, 2015. URL <https://github.com/nbatfai/ram>.
- [15] N. Bátfai. Samu, 2015. URL <https://github.com/nbatfai/samu>.
- [16] N. Bátfai, R. Besenczi, A. Mamenyák, and M. Ispány. OOCWC: The Robocar World Championship Initiative. In *IEEE Conference Publications*, 10.1109/ConTEL.2015.7231223, pages 1–6. Telecommunications (ConTEL), 2015 13th International Conference on, 2015.
- [17] D. Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In *IN PROCEEDINGS OF THE 25TH IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR 2012)*, pages 3642–3649, 2012.
- [18] R. M. French. Catastrophic forgetting in connectionist networks: Causes, consequences and solutions. *Trends in Cognitive Sciences*, 3(4):128–135, 1999.
- [19] K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36:193–202, 1980.
- [20] A. Gosavi. *Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning*. Kluwer Academic Publishers, 2003.
- [21] A. Gosavi. NEURAL NETWORKS AND REINFORCEMENT LEARNING, 2015. URL http://web.mst.edu/~gosavia/neural_networks_RL.pdf.
- [22] L.-J. Lin. *Reinforcement Learning for Robots Using Neural Networks*. PhD thesis, Carnegie Mellon University, 1992. URL <http://www.dtic.mil/dtic/tr/fulltext/u2/a261434.pdf>.
- [23] H. V. Madhyastha, N. Balakrishnan, and K. R. Ramakrishnan. Event information extraction using link grammar. In *13th International Workshop on Research Issues in Data Engineering: Multi-lingual Information Management (RIDE’03)*, pages 16–22. Published by the IEEE Computer Society, 2003.
- [24] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.
- [25] J. D. Oldham. Various Translations of Psalm 23, 2006. URL <http://theory.stanford.edu/~oldham/church/ps23-translations-2006Feb22/ps23-translations-2006Feb22.pdf>.

- [26] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, 3 edition, 2010.
- [27] D. Rusu, L. Dali, B. Fortuna, M. Grobelnik, and D. Mladenic. Triplet Extraction From Sentences. *Proceedings of the 10th International Multi-conference "Information Society - IS 2007"*, A:218–222, 2007.
- [28] S. M. Shieber. Lessons from a Restricted Turing Test. *ArXiv e-prints*, abs/cmp-lg/9404002, 1994. URL <http://arxiv.org/abs/cmp-lg/9404002>.
- [29] A. Stoytchev. Some Basic Principles of Developmental Robotics. *IEEE T. Autonomous Mental Development*, 1(2):122–130, 2009.
- [30] C. Touzet. Neural reinforcement learning for an obstacle avoidance behavior. In *Self Learning Robots, IEE Colloquium on*, page 6/1. IET Conference Publications, 1996.
- [31] T. A. Welch. A technique for high-performance data compression. *Computer*, 17(6):8–19, 1984.